SEQUENCE LISTING

<110> Beraud, Christophe
      Craven, Andrew
      Yu, Ming
      Sakowicz, Roman
      Patel, Umesh A.
      Davies, Katherine A.

<120> NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE

<130> 020552-001410US

<140> US 09/883096
<141> 2001-06-15

<150> US 09/594,655
<151> 2000-06-15

<160> 6

<170> PatentIn Ver. 2.1

<210> 1
<211> 4108
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleic acid sequence of human kinesin motor
      protein gene HsKip3a (Figure 1).

<220>
<223> Description of Artificial Sequence: HsKip3a gene.

<400> 1
```
gcggccgcga attcggcacc aggggcgctc tctcccggtg tgggtactgc tgtctgtggt 60
gtggctgtgg gacccgtgag caagcagcga cgccagcggc ggagaaccga cgaaaggtgt 120
caccacagtg atggcagtgg aggacagcac gctgcaagta gtggtacggg tgcggccccc 180
cacccctcgg gagctggaca gtcagcggcg gccagtggtt caggtggtgg acgagcgggt 240
gctggtgttt aaccctgagg agcccgatgg agggttccct ggcctgaaat ggggtggcac 300
ccatgatggc cccaagaaga agggcaaaga cctgacgttt gtctttgacc gggtctttgg 360
cgaggcggcc acccaacagg acgtgttcca gcacaccacg cacagcgtcc tggacagctt 420
cctccagggc tacaactgct cagtgtttgc ctacggggcc accggggctg ggaagacaca 480
caccatgctg ggaagggagg gggaccccgg catcatgtac ctgaccaccg tggaactgta 540
caggcgcctg gaggcccgcc agcaggagaa gcacttcgag gtgctcatca gctaccagga 600
ggtgtataat gaacagatcc atgacctcct ggagcccaag gggccccttg ccatccgcga 660
ggaccccgac aagggggtgg tggtgcaagg actttctttc caccagccag cctcagccga 720
gcagctgctg gagatactga ccaggggggaa ccgtaaccgc acgcagcacc ccactgatgc 780
caacgcgact tcctcccgct cccatgccat cttccagatc tttgtgaagc agcaggaccg 840
ggttccagga ctgacccagg ctgtccaggt ggccaagatg agcctgattg acctggctgg 900
ctcagagcgg gcatccagca cccatgcgaa gggggagcgg ctgcgggagg gggccaacat 960
caaccgctct ctgctggcgc tcatcaacgt cctcaatgcc ttggccgatg caaagggccg 1020
caagacccat gtgccctacc gggacagcaa actgacccgc ctgctcaaag actccctcgg 1080
gggcaactgc cgcacagtga tgatcgctgc catcagcccc tccagcctga cctacgagga 1140
cacgtacaac accctcaaat atgccgaccg ggccaaggag atcaggctct cgctgaagag 1200
caatgtgacc agcctggact gtcacatcag ccagtatgct accatctgcc aacagctcca 1260
ggctgaggta gccgctctga ggaagaagct ccaagtgtat gagggggggag gccagccccc 1320
accacaggac ctcccaggat ctcccaagtc gggaccacca ccagaacacc ttcccagctc 1380
cccccttgcca ccccacccctc ccagccagcc ctgcacccca gagctccctg cagggcctag 1440
agcccttcaa gaggagagtc tggggatgga ggcccaggtg gagagggcca tggaagggaa 1500
```

```
ctcttcagac caggagcagt ccccagagga tgaggatgaa ggcccagctg aggaggttcc 1560
aacccagatg ccagagcaga accccacaca tgcactgcca gagtcccctc gcctgaccct 1620
gcagcccaag ccagtcgtgg gccacttctc agcacgggaa ctggatgggg accgttctaa 1680
gcagttggcc ctaaaggtgc tgtgcgttgc ccagcggcag tactccctgc tccaagcagc 1740
caacctcctg acgcccgaca tgatcacaga gtttgagacc ctacagcagc tggtgcaaga 1800
ggaaaaaatt gagcctgggg cagaggcctt gaggacttca ggcctggcca gggggggcacc 1860
tctggctcag gagctgtgtt cagagtcaat ccctgtgccg tctcctctct gcccagagcc 1920
tccaggatac actggccctg tgacccggac tatggcgagg cgactgagtg gcccccctgca 1980
caccctggga atcccgcctg acccaactg caccccagcc caggggtccc gatggcccat 2040
ggagaagaag aggaggagac caagcgcctt ggaggcagac agtcccatgg cctcaaagcg 2100
gggcaccaag cgccagcgcc agtccttcct gccctgccta aggagagggt ctctgcctga 2160
cacccaacct tcacaggggc ccagcacccc caaaggagaa agggcctcct cccctgcca 2220
ttcccctcgc gtttgcccag ccacagtcat caaaagccgg gtgcccctgg gcccttccgc 2280
catgcagaac tgctccaccc cgctggctct gcccactcga gacctcaatg ccaccttga 2340
tctctctgag gagcctccct caaagcccag tttccatgaa tgcattggct gggacaaaat 2400
accccaggag ctgagcaggc tggaccagcc cttcatcccc agggcacctg tgcccctgtt 2460
caccatgaag ggccccaagc caacatcttc cctccctggg acctctgcct gcaagaagaa 2520
gcgcgttgcg agttcctcag tctcccatgg ccgcagccgc atcgcccgcc tccccagcag 2580
cactttgaag aggccagctg ggccccttgt actcccagag ctgcccttga gtccctgtg 2640
ccctagcaac cggaggaatg gaaaggacct catcagggtg gggagagcgc tctcagcagg 2700
gaacggcgtc accaaggtgt cctgaccgcc agaatgtcct gaccaccaag gtgtcctaac 2760
ctaccggccc ctctgctgga taccctctt ggacctgtag ccacctgcac caggagctgg 2820
acctgccttc cttacctggg agcaattagt gccaacacac ctttgctgta ttaacatccc 2880
tccccagaca tccatcctgc tactcaccct ctgttaatct cctgttacac tcagcttctt 2940
ggcatgtaca tattcatttg tgagtgttaa tgtgctgctg ttttttgttt tttggtggtt 3000
tttgtttttt gtttttttg ttttgagatg gagtcttact ctgtcgccca ggctggagtg 3060
cagtggtacg atcttggctc actgcaacct ccgcctcctg ggttcaagta attctcctgc 3120
ctcagctttc caagtagctg ggattacagg cacccatcac cacacccagc taattttcgt 3180
ctttttaata gagagggggt ttttccatgt tggccaggct ggtcttgaac tcctgacctc 3240
aggtgatccg cctgcctcag cttcccaaag tgctgagatt acaggcatga gctaccacgc 3300
ctggcccgtg ttgctgtttt aaaggtgctg ccatgttccc ccatctttt ttttttgag 3360
atggagtctc gctctgtcgc ccaggctgga gtgcagtggt ggcgatcttg gctcactgca 3420
agctccgcct cccaggttca caccattctc ctgcctcagc ctcccaagta gctgggacta 3480
caggcgccca ccaccacgcc cggctaattt tttgtatttt tagtagagat ggggtttcac 3540
cgtgttagcc aggctggtct cgatctgacc tcatgatcca cccgcctcgg cctcccaaag 3600
tgctgggatt acaggcgtga gccactgcgc ccggcctccc ctctcattta tgatgccctc 3660
tgtgcaggca gacggctctt gggctctttt ccccacctgt ctctaacaca ggccccacgg 3720
tgatggccac aggcagtaga ggaggaatga ggatgggttg gggagcgggg agtcgcggct 3780
tggctcttcc tggtttctga gagggacatc ttcatcccta ctcccttgg tccccaacca 3840
cagtcctggt gaagatgtgg atgataatgg tgccttgatt ccaaatgaa gacagcttta 3900
ttgcttaact ctattgtaca taggatacac gttcagtgta aaataaagtg taaaggggaa 3960
ttcaggctta atgctgcacc tagatataaa tgctaatgat acttgggttt atagccttct 4020
gatcctttat ttctgcatat atatatagat atatacatat attttggta taacaataaa 4080
ccgtctccat ccttgggaaa aaaaaaaa                                  4108
```

<210> 2
<211> 864
<212> PRT
<213> Artificial Sequence

<220>
<223> Amino acid sequence encoded by human kinesin motor
      protein gene HsKip3a (Figure 1).

<220>
<223> Description of Artificial Sequence:Amino acid
      sequence of HsKip3a.

<400> 2

```
Met Ala Val Glu Asp Ser Thr Leu Gln Val Val Val Arg Val Arg Pro
 1               5                  10                  15

Pro Thr Pro Arg Glu Leu Asp Ser Gln Arg Arg Pro Val Val Gln Val
            20                  25                  30

Val Asp Glu Arg Val Leu Val Phe Asn Pro Glu Glu Pro Asp Gly Gly
        35                  40                  45

Phe Pro Gly Leu Lys Trp Gly Gly Thr His Asp Gly Pro Lys Lys Lys
    50                  55                  60

Gly Lys Asp Leu Thr Phe Val Phe Asp Arg Val Phe Gly Glu Ala Ala
65                  70                  75                  80

Thr Gln Gln Asp Val Phe Gln His Thr Thr His Ser Val Leu Asp Ser
                85                  90                  95

Phe Leu Gln Gly Tyr Asn Cys Ser Val Phe Ala Tyr Gly Ala Thr Gly
            100                 105                 110

Ala Gly Lys Thr His Thr Met Leu Gly Arg Glu Gly Asp Pro Gly Ile
            115                 120                 125

Met Tyr Leu Thr Thr Val Glu Leu Tyr Arg Arg Leu Glu Ala Arg Gln
    130                 135                 140

Gln Glu Lys His Phe Glu Val Leu Ile Ser Tyr Gln Glu Val Tyr Asn
145                 150                 155                 160

Glu Gln Ile His Asp Leu Leu Glu Pro Lys Gly Pro Leu Ala Ile Arg
                165                 170                 175

Glu Asp Pro Asp Lys Gly Val Val Val Gln Gly Leu Ser Phe His Gln
                180                 185                 190

Pro Ala Ser Ala Glu Gln Leu Leu Glu Ile Leu Thr Arg Gly Asn Arg
            195                 200                 205

Asn Arg Thr Gln His Pro Thr Asp Ala Asn Ala Thr Ser Ser Arg Ser
    210                 215                 220

His Ala Ile Phe Gln Ile Phe Val Lys Gln Gln Asp Arg Val Pro Gly
225                 230                 235                 240

Leu Thr Gln Ala Val Gln Val Ala Lys Met Ser Leu Ile Asp Leu Ala
                245                 250                 255

Gly Ser Glu Arg Ala Ser Ser Thr His Ala Lys Gly Glu Arg Leu Arg
            260                 265                 270

Glu Gly Ala Asn Ile Asn Arg Ser Leu Leu Ala Leu Ile Asn Val Leu
            275                 280                 285

Asn Ala Leu Ala Asp Ala Lys Gly Arg Lys Thr His Val Pro Tyr Arg
    290                 295                 300

Asp Ser Lys Leu Thr Arg Leu Leu Lys Asp Ser Leu Gly Gly Asn Cys
305                 310                 315                 320
```

Arg Thr Val Met Ile Ala Ala Ile Ser Pro Ser Ser Leu Thr Tyr Glu
            325   330   335

Asp Thr Tyr Asn Thr Leu Lys Tyr Ala Asp Arg Ala Lys Glu Ile Arg
            340   345   350

Leu Ser Leu Lys Ser Asn Val Thr Ser Leu Asp Cys His Ile Ser Gln
            355   360   365

Tyr Ala Thr Ile Cys Gln Gln Leu Gln Ala Glu Val Ala Ala Leu Arg
    370   375   380

Lys Lys Leu Gln Val Tyr Glu Gly Gly Gly Gln Pro Pro Pro Gln Asp
385   390   395   400

Leu Pro Gly Ser Pro Lys Ser Gly Pro Pro Glu His Leu Pro Ser
            405   410   415

Ser Pro Leu Pro Pro His Pro Pro Ser Gln Pro Cys Thr Pro Glu Leu
            420   425   430

Pro Ala Gly Pro Arg Ala Leu Gln Glu Glu Ser Leu Gly Met Glu Ala
            435   440   445

Gln Val Glu Arg Ala Met Glu Gly Asn Ser Ser Asp Gln Glu Gln Ser
    450   455   460

Pro Glu Asp Glu Asp Glu Gly Pro Ala Glu Glu Val Pro Thr Gln Met
465   470   475   480

Pro Glu Gln Asn Pro Thr His Ala Leu Pro Glu Ser Pro Arg Leu Thr
            485   490   495

Leu Gln Pro Lys Pro Val Val Gly His Phe Ser Ala Arg Glu Leu Asp
            500   505   510

Gly Asp Arg Ser Lys Gln Leu Ala Leu Lys Val Leu Cys Val Ala Gln
    515   520   525

Arg Gln Tyr Ser Leu Leu Gln Ala Ala Asn Leu Leu Thr Pro Asp Met
    530   535   540

Ile Thr Glu Phe Glu Thr Leu Gln Gln Leu Val Gln Glu Glu Lys Ile
545   550   555   560

Glu Pro Gly Ala Glu Ala Leu Arg Thr Ser Gly Leu Ala Arg Gly Ala
            565   570   575

Pro Leu Ala Gln Glu Leu Cys Ser Glu Ser Ile Pro Val Pro Ser Pro
            580   585   590

Leu Cys Pro Glu Pro Pro Gly Tyr Thr Gly Pro Val Thr Arg Thr Met
            595   600   605

Ala Arg Arg Leu Ser Gly Pro Leu His Thr Leu Gly Ile Pro Pro Gly
    610   615   620

Pro Asn Cys Thr Pro Ala Gln Gly Ser Arg Trp Pro Met Glu Lys Lys
625   630   635   640

```
Arg Arg Arg Pro Ser Ala Leu Glu Ala Asp Ser Pro Met Ala Ser Lys
            645                 650             655

Arg Gly Thr Lys Arg Gln Arg Gln Ser Phe Leu Pro Cys Leu Arg Arg
            660                 665             670

Gly Ser Leu Pro Asp Thr Gln Pro Ser Gln Gly Pro Ser Thr Pro Lys
            675                 680             685

Gly Glu Arg Ala Ser Ser Pro Cys His Ser Pro Arg Val Cys Pro Ala
    690                 695             700

Thr Val Ile Lys Ser Arg Val Pro Leu Gly Pro Ser Ala Met Gln Asn
705                 710             715                 720

Cys Ser Thr Pro Leu Ala Leu Pro Thr Arg Asp Leu Asn Ala Thr Phe
            725                 730             735

Asp Leu Ser Glu Glu Pro Pro Ser Lys Pro Ser Phe His Glu Cys Ile
            740                 745             750

Gly Trp Asp Lys Ile Pro Gln Glu Leu Ser Arg Leu Asp Gln Pro Phe
            755                 760             765

Ile Pro Arg Ala Pro Val Pro Leu Phe Thr Met Lys Gly Pro Lys Pro
    770                 775             780

Thr Ser Ser Leu Pro Gly Thr Ser Ala Cys Lys Lys Lys Arg Val Ala
785                 790             795                 800

Ser Ser Ser Val Ser His Gly Arg Ser Arg Ile Ala Arg Leu Pro Ser
            805                 810             815

Ser Thr Leu Lys Arg Pro Ala Gly Pro Leu Val Leu Pro Glu Leu Pro
            820                 825             830

Leu Ser Pro Leu Cys Pro Ser Asn Arg Arg Asn Gly Lys Asp Leu Ile
            835                 840             845

Arg Val Gly Arg Ala Leu Ser Ala Gly Asn Gly Val Thr Lys Val Ser
    850                 855             860
```

<210> 3
<211> 1014
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: HsKip3a
      fragment

<220>
<223> Nucleotide sequence encoding motor domain fragment
      of HsKip3a (Figure 2).

<400> 3
gacagcacgc tgcaagtagt ggtacgggtg cggcccccca cccctcggga gctggacagt 60

```
cagcggcggc cagtggttca ggtggtggac gagcgggtgc tggtgtttaa ccctgaggag 120
cccgatggag ggttccctgg cctgaaatgg ggtggcaccc atgatggccc caagaagaag 180
ggcaaagacc tgacgtttgt ctttgaccgg gtctttggcg aggcggccac ccaacaggac 240
gtgttccagc acaccacgca cagcgtcctg gacagcttcc tccagggcta caactgctca 300
gtgtttgcct acggggccac cggggctggg aagacacaca ccatgctggg aagggagggg 360
gaccccggca tcatgtacct gaccaccgtg gaactgtaca ggcgcctgga ggcccgccag 420
caggagaagc acttcgaggt gctcatcagc taccaggagg tgtataatga acagatccat 480
gacctcctgg agcccaaggg gcccccttgcc atccgcgagg accccgacaa ggggggtggtg 540
gtgcaaggac tttcttcca ccagccagcc tcagccgagc agctgctgga gatactgacc 600
aggggggaacc gtaaccgcac gcagcacccc actgatgcca acgcgacttc ctcccgctcc 660
catgccatct tccagatctt tgtgaagcag caggaccggg ttccaggact gacccaggct 720
gtccaggtgg ccaagatgag cctgattgac ctggctggct cagagcgggc atccagcacc 780
catgcgaagg gggagcggct gcgggagggg gccaacatca accgctctct gctggcgctc 840
atcaacgtcc tcaatgcctt ggccgatgca aagggccgca gacccatgt gccctaccgg 900
gacagcaaac tgacccgcct gctcaaagac tccctcgggg gcaactgccg cacagtgatg 960
atcgctgcca tcagcccctc cagcctgacc tacgaggaca cgtacaacac cctc      1014
```

```
<210> 4
<211> 338
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: HsKip3a fragment.

<220>
<223> Amino acid sequence of the motor domain fragment
      of HsKip3a (Figure 3).

<400> 4
```

```
Asp Ser Thr Leu Gln Val Val Val Arg Val Arg Pro Pro Thr Pro Arg
 1           5               10              15

Glu Leu Asp Ser Gln Arg Arg Pro Val Val Gln Val Val Asp Glu Arg
            20              25              30

Val Leu Val Phe Asn Pro Glu Glu Pro Asp Gly Gly Phe Pro Gly Leu
         35              40              45

Lys Trp Gly Gly Thr His Asp Gly Pro Lys Lys Gly Lys Asp Leu
    50              55              60

Thr Phe Val Phe Asp Arg Val Phe Gly Glu Ala Ala Thr Gln Gln Asp
 65              70              75              80

Val Phe Gln His Thr Thr His Ser Val Leu Asp Ser Phe Leu Gln Gly
            85              90              95

Tyr Asn Cys Ser Val Phe Ala Tyr Gly Ala Thr Gly Ala Gly Lys Thr
            100             105             110

His Thr Met Leu Gly Arg Glu Gly Asp Pro Gly Ile Met Tyr Leu Thr
        115             120             125

Thr Val Glu Leu Tyr Arg Arg Leu Glu Ala Arg Gln Gln Glu Lys His
    130             135             140

Phe Glu Val Leu Ile Ser Tyr Gln Glu Val Tyr Asn Glu Gln Ile His
145             150             155             160
```

Asp Leu Leu Glu Pro Lys Gly Pro Leu Ala Ile Arg Glu Asp Pro Asp
                165   170   175

Lys Gly Val Val Val Gln Gly Leu Ser Phe His Gln Pro Ala Ser Ala
            180   185   190

Glu Gln Leu Leu Glu Ile Leu Thr Arg Gly Asn Arg Asn Arg Thr Gln
        195   200   205

His Pro Thr Asp Ala Asn Ala Thr Ser Ser Arg Ser His Ala Ile Phe
    210   215   220

Gln Ile Phe Val Lys Gln Gln Asp Arg Val Pro Gly Leu Thr Gln Ala
225   230   235   240

Val Gln Val Ala Lys Met Ser Leu Ile Asp Leu Ala Gly Ser Glu Arg
            245   250   255

Ala Ser Ser Thr His Ala Lys Gly Glu Arg Leu Arg Glu Gly Ala Asn
            260   265   270

Ile Asn Arg Ser Leu Leu Ala Leu Ile Asn Val Leu Asn Ala Leu Ala
        275   280   285

Asp Ala Lys Gly Arg Lys Thr His Val Pro Tyr Arg Asp Ser Lys Leu
    290   295   300

Thr Arg Leu Leu Lys Asp Ser Leu Gly Gly Asn Cys Arg Thr Val Met
305   310   315   320

Ile Ala Ala Ile Ser Pro Ser Ser Leu Thr Tyr Glu Asp Thr Tyr Asn
            325   330   335

Thr Leu

<210> 5
<211> 383
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: HsKip3a
      fragment

<220>
<223> Amino acid sequence of HsKip3a fragment used in
      the ATPase assay (Figure 4).

<400> 5
Met Ala Val Glu Asp Ser Thr Leu Gln Val Val Val Arg Val Arg Pro
  1   5   10   15

Pro Thr Pro Arg Glu Leu Asp Ser Gln Arg Arg Pro Val Val Gln Val
            20   25   30

Val Asp Glu Arg Val Leu Val Phe Asn Pro Glu Glu Pro Asp Gly Gly
        35   40   45

Phe Pro Gly Leu Lys Trp Gly Gly Thr His Asp Gly Pro Lys Lys Lys
    50   55   60

Gly Lys Asp Leu Thr Phe Val Phe Asp Arg Val Phe Gly Glu Ala Ala
 65   70   75   80

Thr Gln Gln Asp Val Phe Gln His Thr Thr His Ser Val Leu Asp Ser
                85   90   95

Phe Leu Gln Gly Tyr Asn Cys Ser Val Phe Ala Tyr Gly Ala Thr Gly
                100   105   110

Ala Gly Lys Thr His Thr Met Leu Gly Arg Glu Gly Asp Pro Gly Ile
            115   120   125

Met Tyr Leu Thr Thr Val Glu Leu Tyr Arg Arg Leu Glu Ala Arg Gln
    130   135   140

Gln Glu Lys His Phe Glu Val Leu Ile Ser Tyr Gln Glu Val Tyr Asn
145   150   155   160

Glu Gln Ile His Asp Leu Leu Glu Pro Lys Gly Pro Leu Ala Ile Arg
                165   170   175

Glu Asp Pro Asp Lys Gly Val Val Val Gln Gly Leu Ser Phe His Gln
                180   185   190

Pro Ala Ser Ala Glu Gln Leu Leu Glu Ile Leu Thr Arg Gly Asn Arg
            195   200   205

Asn Arg Thr Gln His Pro Thr Asp Ala Asn Ala Thr Ser Ser Arg Ser
    210   215   220

His Ala Ile Phe Gln Ile Phe Val Lys Gln Gln Asp Arg Val Pro Gly
225   230   235   240

Leu Thr Gln Ala Val Gln Val Ala Lys Met Ser Leu Ile Asp Leu Ala
                245   250   255

Gly Ser Glu Arg Ala Ser Ser Thr His Ala Lys Gly Glu Arg Leu Arg
            260   265   270

Glu Gly Ala Asn Ile Asn Arg Ser Leu Leu Ala Leu Ile Asn Val Leu
            275   280   285

Asn Ala Leu Ala Asp Ala Lys Gly Arg Lys Thr His Val Pro Tyr Arg
    290   295   300

Asp Ser Lys Leu Thr Arg Leu Leu Lys Asp Ser Leu Gly Gly Asn Cys
305   310   315   320

Arg Thr Val Met Ile Ala Ala Ile Ser Pro Ser Ser Leu Thr Tyr Glu
                325   330   335

Asp Thr Tyr Asn Thr Leu Lys Tyr Ala Asp Arg Ala Lys Glu Ile Arg
                340   345   350

Leu Lys Gly Asn Ser Lys Leu Glu Gly Lys Pro Ile Pro Asn Pro Leu
    355   360   365

Leu Gly Leu Asp Ser Thr Arg Thr Gly His His His His His His
    370   375   380

<210> 6
<211> 1152
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: HsKip3a
     fragment

<220>
<223> Nucleotide sequence of HsKip3a fragment used in
     ATPase assay (Figure 5).

<400> 6
```
atggcagtgg aggacagcac gctgcaagta gtggtacggg tgcggccccc cacccctcgg 60
gagctggaca gtcagcggcg gccagtggtt caggtggtgg acgagcgggt gctggtgttt 120
aaccctgagg agcccgatgg agggttccct ggcctgaaat ggggtggcac ccatgatggc 180
cccaagaaga agggcaaaga cctgacgttt gtctttgacc gggtctttgg cgaggcggcc 240
acccaacagg acgtgttcca gcacaccacg cacagcgtcc tggacagctt cctccagggc 300
tacaactgct cagtgtttgc ctacggggcc accggggctg ggaagacaca caccatgctg 360
ggaagggagg gggacccggg catcatgtac ctgaccaccg tggaactgta caggcgcctg 420
gaggcccgcc agcaggagaa gcacttcgag gtgctcatca gctaccagga ggtgtataat 480
gaacagatcc atgacctcct ggagcccaag gggcccccttg ccatccgcga ggaccccgac 540
aaggggggtgg tggtgcaagg actttctttc caccagccag cctcagccga gcagctgctg 600
gagatactga ccaggggggaa ccgtaaccgc acgcagcacc ccactgatgc caacgcgact 660
tcctcccgct cccatgccat cttccagatc tttgtgaagc agcaggaccg ggttccagga 720
ctgacccagg ctgtccaggt ggccaagatg agcctgattg acctggctgg ctcagagcgg 780
gcatccagca cccatgcgaa ggggggagcgg ctgcgggagg gggccaacat caaccgctct 840
ctgctggcgc tcatcaacgt cctcaatgcc ttggccgatg caaagggccg caagacccat 900
gtgccctacc gggacagcaa actgacccgc ctgctcaaag actccctcgg gggcaactgc 960
cgcacagtga tgatcgctgc catcagcccc tccagcctga cctacgagga cacgtacaac 1020
accctcaaat atgccgaccg ggccaaggag atcaggctca agggcaattc gaagcttgaa 1080
ggtaagccta tccctaaccc tctcctcggt ctcgattcta cgcgtaccgg tcatcatcac 1140
catcaccatt ga                                                   1152
```